

Rank aggregation methods dealing with incomplete information applied to smart cities

E. Dopazo
School of Computer Engineering
Technical University of Madrid
Madrid, Spain 28660
Email: edopazo@fi.upm.es

M.L. Martínez-Céspedes
School of Computer Engineering
Technical University of Madrid
Madrid, Spain 28660
Email: mmartinez@a-cing.com

Fuzzy decision-making: Consensus and Missing preferences

Abstract—City-rankings have become a central instrument for assessing the attractiveness of urban regions over the last years. Demographic, environmental, economic, political and socio-cultural factors are forcing the urban world to design and implement Smart Cities. A set of multidimensional components underlies the fuzzy smart city concept. As a result cities are evaluated and ranked with regard to different characteristics, and smart city measures are achieved through chosen indicators. Therefore, the problem of combining multiple rankings to form an aggregate ranking, which compares city performance, is recognized as a useful tool in this context. Moreover, a usual situation is when incomplete information arises and only partial rankings may be supplied. This paper addresses the general problem of rank aggregation dealing with incomplete information based on rank aggregation methods and multicriteria decision making theory. It consists on constructing a consensus ranking from partial rankings of a set of objects provided according different criteria. Our techniques rely on outranking matrices as a way of collecting relevance information from input data, theory of fuzzy preference relations and the PageRank algorithm.

Index Terms—Rank Aggregation, Incomplete Information, Fuzzy Preference Relation, Smart City.

I. INTRODUCTION

The Smart City is of great complexity and with a multicriteria inherent character. City-rankings have become a central instrument for assessing the attractiveness of urban regions over the last 20 years ([12],[22]). Demographic, environmental, economic, political and socio-cultural factors are forcing the urban world to design and implement Smart Cities. There is no agreement on the exact definition of the smart city concept. It can consider that a smart city represents a unified urban entity according different criteria. The main dimensions of a Smart City can be identified through literature review and it includes: smart economy, smart mobility, smart environment, sustainable city, smart people, smart living, and smart governance, digital city, etc. ([4], [6] and [12]).

The definition of metrics in the field of smart cities is driven by two principles: First, to compare cities between themselves and learn from the best, and second to understand the internal dynamics of smart cities, define weaknesses, and recognize the effort needed to overcome them. The smart city measures are achieved through carefully chosen indicators and allow cities to reorganize itself successfully, via an understanding

of its strengths and weaknesses. In these kinds of comparative studies, cities are evaluated and ranked with regard to different economic, social, environmental, and geographical characteristics in order to reveal the best places for certain activities. The number of rankings in existence around the world is growing year after year. Therefore, the problem of combining multiple rankings to form a compromise or consensus ranking, which compare city performance, is recognized as a useful tool in policy analysis, city marketing, benchmarking and public communication.

Moreover, a usual situation in this context is when information of all items is not possible and only partial rankings [22] (concerning a subset of the studied cities) may be supplied. According to these considerations, a challenge in this context is to provide methods and algorithms to generate an aggregate Smart City Ranking to be able to deal with limited, conflicting and heterogeneous information from different sources. Then, we focus on the general problem of constructing a complete ranking of a set of candidates that best represents, in a precise sense, possibly conflicting and incomplete information given by regards partial rankings according multiple criteria.

The problem of rank aggregation ([12],[23]) is not new and appears in many applications. There has been an extensive body of work on this topic, beginning with the works of Borda and Condorcet in Social Choice Theory (s. XVIII). Kemeny-Snell [18] proposed to determine the best aggregate ranking that minimizes the sum of the deviations from the individual rankings (in terms of a distance between rankings) However this optimization problem is shown to be NP-hard [9]. Cook and Seiford [5] provides a consensus ranking in terms of the median ranking. A goal programming method dealing with interval rankings in a group decision-making problem is studied in [13]. Recently, some methods have been applied in the context of meta-search engines ([9], [25]) and information retrieval [10], among others.

A great deal of the literature, however is concerned with models that assume full lists. These models can not be used in the context of smart cities, where most of the indicators provide partial lists that evaluate and compare only proper subsets of the candidates.

In this scenario, we propose methods and algorithms for ranking the candidates under consideration, by computing

priority vectors that best reflect limited information given across partial rankings. Our techniques are articulated in two stages. In the first stage, a multicriteria decision method is used to obtain the weights of considered indicators or evaluation criteria by a panel of stakeholders or experts, in the case they have not been previously stated. In the second stage (our main concern) we provide two methods to determine a complete aggregate ranking. They are based on deriving priority vectors of the candidates from outranking matrices that summarize preferences from partial lists, in some precise sense. In the first method, we use theory of fuzzy preference relations [27]. In the second method, a procedure similar to Google PageRank algorithm [3] is proposed.

The rest of the paper is organized as follows. Section II is devoted to establish the smart city context and to formulate the general problem we address. Main definitions and notation are also introduced. In section III, we propose rank aggregation methods dealing with incomplete information. An illustrating numerical example is included in Section IV. Finally, main conclusions derived from this work are given in Section V.

II. MEASURING THE SMART CITY ISSUE. PROBLEM FORMULATION

On the other side, cities are engines of innovation, economic growth, and technological progress. Cities must lead the way to show how our society can be greener, more intelligent, more efficient, more attractive, more social, more digital, and give everyone a better quality of life. The distinctive quality of any city is its capacity to aggregate people, skills, talents, resources, capital, businesses, and government, social and physical infrastructure. This “coming together” is supported by complex systems for land use, transportation, utilities, sanitation, housing, and public services.

Although there is no agreement on the exact definition of the smart city concept, a number of main dimensions of a Smart City can be identified through literature review ([12], [6] and [4]). It can consider that a smart city represents a unified urban entity which integrates three types of city in one [1], which are:

- 1) *The Digital City*: The terms Digital Community, Digital City or Information City and e-city are also used. It refers to a connected community that combines broadband communications infrastructure; flexible service-oriented computing infrastructure based on open industry standards; and innovative services to meet the needs of governments and their employees, citizens and businesses.
- 2) *The Sustainable City*: It refers to a city designed with consideration of environmental impact, inhabited by people encouraged to minimize the required inputs of energy, water and food, and waste output of heat, air pollution and water pollution. It represents a kind of resilient city.
- 3) *The Knowledge City*: It refers to an innovative and competitive city with: innovative industries, clusters of industries by sector, districts of a city, knowledge

workforce as education and employment, and able to create knowledge-intensive companies.

Therefore, we conclude that the Smart City concept is of great complexity and with a multi-criteria inherent character. There is a lot of research on selecting the adequate metrics [21] to measure the Smart City concept, also in the metadata structure [11], but there is a lack of research in how to aggregate this information through an established model. Our work follows this line of research. More specifically, this work extends previous studies ([23], [24]) and presents a methodology to construct a consensus ranking in virtue of the idea of rank aggregation methods using available data. The problem of rank a set of cities attending its smart city nature can be considered a multi-criteria problem, due to the multidimensionality character of the Smart City concept. Moreover, a usual situation in this context is when information of all cities is not possible because large number of cities or incomplete information, this is, each indicator provides a partial list that evaluates and compares only a proper subset of the candidates.

A. Preliminaries. Problem formulation

We next introduce some basic definitions (see [9]) and notation to formulate rigorously the general problem we address. Given a set of n cities (candidates, alternatives) $X = \{x_1, \dots, x_n\}$, an ordered list (or simply ranking) R with respect to X is an ordering (or permutation) of a subset S of X , i.e., $R = \{x_{i_1} \geq x_{i_2} \geq \dots x_{i_l}\}$, where each $x_{i_k} \in S$, and \geq represents some ordering relation on S . Also, if $x_i \in X$ is present in R , r_i denotes the position or rank of the element x_i in ranking R . Without loss of generality, we can assume that a highly ranked or preferred element has a low-numbered position in the list.

According to the elements ranked in the list R , some situations arise: If R contains all the elements of X , then it is said to be a *full list*, that means it is a total ordering of X . In many situations full lists are not possible. In this case, where only the elements of a proper subset S of X are ranked, R is called *partial list*. A special case of partial lists is where the subset S of ordered elements corresponds to the top k ranked elements. The elements that are not included in the list are assumed to be ranked below k position. Such lists are called *top k rank lists*, where k is the size of the list. In order to refer all these situations in a uniform way, we will represent a rank ordering R by the ranking vector $R = (r_i)$, where r_i denotes the position or rank of element x_i in this ranking, otherwise $r_i = \star$ indicates x_i is not present in the list (there is no information about this candidate).

Now we are in conditions to formulate the general problem we address. Let $X = \{x_1, \dots, x_n\}$ be a set of n elements (cities, candidates) that have ranked attending a multicriteria concept that it is articulated through a variety of m criteria, $m > 1$. We consider the evaluation of the elements with respect to an individual criterion or indicator k is given by a partial ranking R^k on X , where $k = 1, \dots, m$. Let us also assume the existence of a set of individual indicator

weights $w = (w_k)$, $k = 1, \dots, m$, with $\sum_{k=1}^m w_k = 1$, derived as importance coefficients. Then, the problem to be dealt with is to construct a complete consensus ranking on X from available ordinal information (contained in the input rankings R^k , $k = 1, \dots, m$), which is possibly incomplete and conflicting.

III. RANK AGGREGATION METHODS DEALING WITH PARTIAL LISTS

Let $X = \{x_1, x_2, \dots, x_n\}$ ($n \geq 2$) be the set of candidates. Given m (full or partial) ordered lists or rankings $R^k = (r_i^k)$, where r_i^k denotes the position or rank of alternative x_i respect to k th criterion, otherwise $r_i^k = \star$ indicates there is no information about x_i with respect to this criterion. Our objective is to produce an aggregate ranking of the candidates that summarize as best as possibly, in some precise sense, input information.

In order to accomplish this goal, first, we assume the existence of a set of importance weights (w_1, \dots, w_m) that represents the relative importance assigned of the criteria (rankings), where $w_k > 0$ and $\sum_{k=1}^m w_k = 1$, given by a group of experts or stakeholders. In the case these weights have not been stated a priori, we can consider the AHP (Analytic Hierarchy Process) [26], a leading multi-criteria decision making method, to guide to the group of experts to derive such weights. In AHP, the decision problem is structured as a hierarchy of criteria and sub-criteria. Then the experts express their judgments respect to an objective by regards a pairwise comparison matrix $W = (w_{ij})$. The entry w_{ij} of the matrix reflects the importance ratio between criterion i and criteria j according to the Saaty scale. Then the relative weights $w = (w_1, \dots, w_m)$ of the rankings or evaluation criteria are calculated by using the eigenvector method [26] or a distance-based method [8].

Once the weight vector w is known, we address the rank aggregation problem dealing with partial rankings (our main concern). We propose two methods, the first one is based on a fuzzy preference relation approach, meanwhile the second one is inspired on the PageRank algorithm. Both methods are articulated as follows:

- 1) We create an aggregate dominance matrix from ranking information. It provides a model for collecting information on ordinal preferences when a direct comparison of all the objects is not possible and only partial rankings are supplied.
- 2) A complete ranking is determined through a priority vector calculated from the previous matrix, according to the corresponding model. Then the objects are ranked in decreasing order attending to the priority vector.

First, in order to handle unknown values in the partial lists $R^k = (r_i^k)$, we construct the dominance matrix $X^k = (x_{ij}^k)$ associated to k th criterion, defined as follows:

$$x_{ij}^k = \begin{cases} 1 & \text{if } r_i^k \neq \star, \quad r_j^k \neq \star, \text{ and } r_i^k < r_j^k \\ 1/2 & \text{if } r_i^k \neq \star, \quad r_j^k \neq \star, \text{ and } r_i^k = r_j^k \\ 0 & \text{if } r_i^k \neq \star, \quad r_j^k \neq \star, \text{ and } r_i^k > r_j^k \\ \star & \text{otherwise,} \end{cases} \quad (1)$$

where \star means no information is available related to the relative position between element x_i and element x_j respect to k th criterion. We will assume $x_{ii}^k = 1/2$, $k = 1, \dots, m$. We note that in the special case R^k be a top d -list ($d < n$), we can consider that:

$$\begin{aligned} x_{ij}^k &= 1 && \text{if } r_i^k \neq \star \text{ and } r_j^k = \star, \\ x_{ij}^k &= 0 && \text{if } r_i^k = \star \text{ and } r_j^k \neq \star, \\ x_{ij}^k &= \star && \text{if } r_i^k \neq \star \text{ and } r_j^k \neq \star. \end{aligned}$$

A. Rank Aggregation Method 1: Fuzzy preference relation approach

We create a $n \times n$ aggregate preference matrix $A = (a_{ij})$, as a model for collecting partial ordinal preferences across R^k , $k = 1, \dots, m$, as follows:

$$a_{ij} = \begin{cases} \sum_{k: x_{ij}^k > 0} w_k x_{ij}^k / \sum_{k: x_{ij}^k \neq \star} w_k & \text{if } \exists k: x_{ij}^k \neq \star, \\ \star & \text{otherwise,} \end{cases} \quad (2)$$

where $i, j = 1, \dots, n$. It quantifies the extent x_i is as good or better than x_j according the available information.

We notice A is an incomplete matrix, where $a_{ij} \neq \star$ indicates the “relative degree or intensity of preference” of item x_i over x_j . In this case, it easy to verify that $a_{ij} \in [0, 1]$. Then, $a_{ij} = 0$ indicates that x_j is absolutely preferred to x_i ; $a_{ij} = 1/2$ means indifference regards items x_i and x_j ; $a_{ij} > 1/2$ indicates that x_i is definitely preferred to x_j and the larger value of a_{ij} , the greater the preference degree of the candidate x_i over x_j . Otherwise, $a_{ij} = \star$ means there is no direct pairwise information about x_i and x_j through the considered criteria (from the input partial lists). A similar matrix has been considered in [23] and [24] for full ranking lists.

We can consider that matrix A describes an incomplete fuzzy preference relation ([14], [15], [27]) on X , that estimates the partial preference information contained in R^k , $k = 1, \dots, m$. Moreover, the binary relation defined in A satisfies the reciprocity property for known values: $a_{ij} + a_{ji} = 1$ if $a_{ij} \neq \star \wedge a_{ji} \neq \star$, $i, j = 1, \dots, n$.

The preference matrix A can be interpreted as an incomplete weighted digraph G , where X is the set of vertices and a_{ij} represents the weight of edge from vertex x_i to vertex x_j if $a_{ij} \neq \star$. There is an edge from a vertex x_i to a vertex x_j in G , if x_j is dominated or equal ranked to x_i for at least one criteria k , $k = 1, \dots, m$. This is, there is an arc (x_i, x_j) in G if there exists k such that $r_i^k \leq r_j^k$, where $r_i^k \neq \star$ and $r_j^k \neq \star$. We will assume that G is a strongly connected digraph, in the sense for each pair of elements (x_i, x_j) , $x_i, x_j \in X$, there exists at least one sequence of arcs $(x_i, x_{i_1}), \dots, (x_{i_k}, x_j)$ in G . It means, that for each pair of different elements, at least a paired comparison is available, directly from input information in R^k , $k = 1, \dots, m$, or indirectly by means of sequence of edges in G . It ensures there is enough information across the input rankings such that it is possible to compare any pair of elements, even if they not have been compared directly

respect to one criterion or indicator. This assumption covers the called total ignorance situation considered in the context of group decision making problem in [2], where one expert of the group is allowed not to provide any preference information involving one alternative.

Once the matrix A has been created, our goal is to derive a ranking of the n items from aggregate preference information in matrix A . Since matrix A does not necessarily satisfy transitive property (some directed cycles may exist), there is not necessarily a permutation on X associated to it. Then the problem could be formulated to find a transitive matrix $B = (b_{ij})$, where $b_{ij} \in \{0, 1/2, 1\}$ (and then the associated rank vector r), such that

$$\|A - B\| \quad (3)$$

is minimal for some matrix norm $\|\cdot\|$. This is a NP-problem ([5], [16], [19]).

To accomplish this, we first focus on deriving a nonnegative priority vector $v = (v_1, \dots, v_m)$, where v_i reflects the priority degree of object x_i from the incomplete fuzzy preference relation expressed in matrix A . In the ideal case, where A is a complete preference relation verifying transitive requirements, there exists a priority vector v such that

$$a_{ij} = \frac{v_i - v_j}{2} + \frac{1}{2}, i, j = 1, \dots, n, \quad (4)$$

where $|v_i - v_j| \leq 1, i, j = 1, \dots, n$ ([27], [15]). Taking into consideration this result, we formulate the problem as finding a nonnegative vector $v = (v_i)$, such that best approximates known data in matrix A attending to (4). This is, we look for a transitive matrix $V = (\frac{1}{2}(v_i - v_j + 1))$ that approximates as best as possibly the incomplete matrix A ,

$$\|A - V\|, \quad (5)$$

in some matrix norm $\|\cdot\|$. The matrix approximation approach has been considered for fuzzy pairwise comparison matrices in [7]. In the case we consider the usual l_2 norm, the optimization problem (5) becomes:

$$\sum_{i,j=1, a_{ij} \neq \star}^n |2a_{ij} - (v_i - v_j) - 1|^2, \quad (6)$$

subject to $v_i \geq 0$ and $|v_i - v_j| \leq 1, i, j = 1, \dots, n$. We note that the proposed model ((5)-(6)) addresses simultaneously the issues of transitivity, incomplete and conflicting information. The problem (6) results into a least squares problem for which many numerical tools are available to compute the solutions.

In the case we consider the l_1 norm in (5), the optimization problem (5) becomes:

$$\sum_{i,j=1, a_{ij} \neq \star}^n |2a_{ij} - (v_i - v_j) - 1|, \quad (7)$$

subject to $v_i \geq 0$ and $|v_i - v_j| \leq 1, i, j = 1, \dots, n$. In order to solve this optimization problem (7), we introduce the new nonnegative variables n_{ij} and p_{ij} , that represent the negative

and positive deviations, respectively, from the known data $a_{ij} \neq \star$, as follows:

$$\begin{aligned} n_{ij} &= \frac{1}{2} \left[|v_i - v_j - 2a_{ij} + 1| + (v_i - v_j - 2a_{ij} + 1) \right], \\ p_{ij} &= \frac{1}{2} \left[|v_i - v_j - 2a_{ij} + 1| - (v_i - v_j - 2a_{ij} + 1) \right], \end{aligned} \quad (8)$$

for $i, j = 1, \dots, n$ such that $a_{ij} \neq \star$. Then, the problem (7) can be formulated in terms of the new variables as follows:

$$\min \sum_{i,j=1, a_{ij} \neq \star}^n (n_{ij} + p_{ij}), \quad (9)$$

subject to:

$$\begin{aligned} v_i - v_j - 2a_{ij} + 1 - n_{ij} + p_{ij} &= 0, i, j = 1, \dots, n : a_{ij} \neq \star \\ n_{ij}, p_{ij} &\geq 0, \quad i, j = 1, \dots, n : a_{ij} \neq \star, \\ v_i &\geq 0, \quad |v_i - v_j| \leq 1, \quad i, j = 1, \dots, n. \end{aligned} \quad (10)$$

It results a linear programming problem that can be solved by direct application of the simplex method.

Once vector $v = (v_i)$ has been computed, the final ranking $R = (r_i)$ of items $x_i, i = 1, \dots, n$, is produced by sorting them in descending order attending to the values $v_i, i = 1, \dots, n$. Then r_i indicates the position or rank of x_i . Ties in the ranking can happen if equal values of v_i are obtained.

Most of the methods dealing with incomplete fuzzy relations in the literature are articulated in several steps, one of them should be a completion phase, where the unknown elements are replaced by numerical values following different strategies [2], [14], in [29] a complete review is included.

B. Rank Aggregation Method 2: Perron eigenvector approach

This second method follows the ideas of Keener's rating method ([17], [19]) and the PageRank algorithm of Brin and Page ([3]). First we build an aggregate dominance matrix $A = (a_{ij})$, collecting the relative comparisons of the strength of the candidates in the following way:

$$a_{ij} = \begin{cases} \frac{\sum_{k: x_{ij}^k > 0} w_k x_{ij}^k}{\sum_{l=1}^n \sum_{k: x_{lj}^k > 0} w_k x_{lj}^k} & \text{if } \exists k: x_{ij}^k \neq \star, \\ 0 & \text{otherwise,} \end{cases} \quad (11)$$

where x_{ij}^k as in (1), and $i, j = 1, \dots, n$. It represents how well the item x_i performs against item x_j . We notice matrix A is a sub-stochastic matrix: $a_{ij} \geq 0$ and $\sum_{i=1}^n a_{ij} \leq 1$, even more, this sum is equal 1 but also the case where $a_{ij} = 0, \forall i$, that corresponds the existence of an undefeated element along all the criteria.

Now, we look for a nonnegative rating vector $r = (r_1, \dots, r_n)^t$ ($\sum_{i=1}^n r_i = 1$), where r_i indicates the strength of x_i , from matrix A . Following the ideas of rating method of Keener and the PageRank algorithm ([17], [19], [3]), the score s_i of element x_i is defined by

$$s_i = \sum_{j=1}^n a_{ij} r_j, \quad i = 1, \dots, n. \quad (12)$$

This is, the score of an element is estimated from their interactions with the other elements along the criteria (the partial rankings $R^k, k = 1, \dots, m$) together the strength of the “opponent element”. Now the strength of the element is presumed to be proportional to its score, that is,

$$s_i = \sum_{j=1}^n a_{ij} r_j = \lambda r_i, \quad i = 1, \dots, n, \quad (13)$$

$$Ar = \lambda r. \quad (14)$$

Then the rating vector r is given by a positive eigenvector associated to the Perron eigenvalue of matrix A , $Ar = \lambda r$ ([17]). In practice the power method, a simple iterative algorithm (basically, $Ar^{k+1} = \lambda r^k$) can be used to obtain an approximation of the rating vector r . Some adjustments have to be made to the matrix A that guarantee the existence of r . As in the previous aggregation method, the directed graph defined by the matrix A is assumed strongly connected, or which it is the same, matrix A is assumed irreducible.

IV. AN ILLUSTRATIVE CASE STUDY

This section illustrates concepts and methods of the previous sections. We consider a simple illustrative example where five cities {Amsterdam, Barcelona, Berlin, Copenhagen, Oslo} have to be ranked according to a composite smart city indicator. Let us assume that three dimensions [1] are considered, i.e. the digital city, the sustainable city and the knowledge city, all of them with the same weight, that is $w_k = 1/3, k = 1, 2, 3$. Now, the input information describing the considered factors of a smart city are derived from public and freely available data, obtained from the following indicators: the *Technology Score Indicator* [20] to evaluate the digital dimension, the *Green City Index* [28] to evaluate the sustainability dimension, and the *Talent Score* [20] to evaluate the knowledge dimension. The Technology Score Indicator is one of the dimensions of the Global Cities Index developed by the Martin Prosperity Institute (<http://martinprosperity.org/>), which benchmarks city performance on the three Ts of economic development: Talent, Technology and Tolerance. In addition the survey also tracks a fourth dimension, the Quality of Place. The Green City Index compares major cities in Europe in terms of their environmental performance and policies. To aid understanding of the strengths and weaknesses of each city and their performance against peers and to foster best practice sharing. As restrictions from the previous mentioned indicators to the undertaken sample of cities, we get the partial rankings in the following table:

Now, the dominance matrices according each one of these indicators are constructed following (1):

$$X^1 = \begin{pmatrix} 0.5 & 1. & \star & 0. & 0. \\ 0. & 0.5 & \star & 0. & 0. \\ \star & \star & 0.5 & \star & \star \\ 1. & 1. & \star & 0.5 & 0.5 \\ 1. & 1. & \star & 0.5 & 0.5 \end{pmatrix},$$

TABLE I
SAMPLE OF CITIES RANKS

Cities	Technology Rank	Green City Rank	Talent Rank
Amsterdam	2	3	2
Barcelona	3	*	3
Berlin	*	4	*
Copenhagen	1	1	2
Oslo	1	2	1

$$X^2 = \begin{pmatrix} 0.5 & \star & 1. & 0. & 0. \\ \star & 0.5 & \star & \star & \star \\ 0. & \star & 0.5 & 0. & 0. \\ 1. & \star & 1. & 0.5 & 1. \\ 1. & \star & 1. & 0. & 0.5 \end{pmatrix},$$

$$X^3 = \begin{pmatrix} 0.5 & 1. & \star & 0.5 & 0. \\ 0. & 0.5 & \star & 0. & 0. \\ \star & \star & 0.5 & \star & \star \\ 0.5 & 1. & \star & 0.5 & 0. \\ 1. & 1. & \star & 1. & 0.5 \end{pmatrix},$$

where 1 indicates i dominates j , 0.5 means there is a tie between i and j , 0 indicates i is dominated by j , and \star means no data is available, along the considered indicator.

Following the procedure of Method 1, described in Subsection III-A, first we create the aggregate preference matrix A_1 from the individual preference relations X^1, X^2 and X^3 :

$$A_1 = \begin{pmatrix} 0.5 & 1. & 1. & 0.1667 & 0. \\ 0. & 0.5 & \star & 0. & 0. \\ 0. & \star & 0.5 & 0. & 0. \\ 0.8333 & 1. & 1. & 0.5 & 0.5 \\ 1. & 1. & 1. & 0.5 & 0.5 \end{pmatrix}.$$

We notice A_1 represents an incomplete fuzzy preference relation, where $a_{ij} \in [0, 1]$ indicates the “relative strength of preference” of city x_i over x_j , otherwise $a_{ij} = \star$ means there is no direct pairwise information about x_i and x_j through the considered rankings. Moreover, the relation defined in A_1 satisfies the reciprocity property for known values: $a_{ij} + a_{ji} = 1$ if $a_{ij} \neq \star \wedge a_{ji} \neq \star, i, j = 1, \dots, 5$.

Now, by solving the resulting least squares problem (6), using MATLAB numerical software, we obtain the priority vector $v = (0.6667, 0., 0., 1.1333, 1.2000)$ from fuzzy preference relation A_1 . It produces the ordering

$$\text{Oslo} > \text{Copenhagen} > \text{Amsterdam} > \text{Barcelona} = \text{Berlin}.$$

To apply the Method 2 proposed in Subsection III-B, first the dominance matrix A_2 is constructed as described in (11):

$$A_2 = \begin{pmatrix} 0.2143 & 0.2667 & 0.2222 & 0.1429 & 0 \\ 0 & 0.2000 & 0 & 0 & 0 \\ 0 & 0 & 0.3333 & 0 & 0 \\ 0.3571 & 0.2667 & 0.2222 & 0.4286 & 0.5000 \\ 0.4286 & 0.2667 & 0.2222 & 0.4286 & 0.5000 \end{pmatrix}.$$

We notice A_2 is an stochastic matrix. Now we compute the rating vector $r = (r_1, \dots, r_5)^t$ as the normalized positive

eigenvector associated to the dominant eigenvalue ($\lambda = 1$) of matrix A_2 . Using MATLAB numerical software, we obtain

$$r = (0.0828, 0., 0., 0.4556, 0.4615)^t,$$

that provides the ordering:

Oslo > Copenhagen > Amsterdam > Barcelona = Berlin.

We conclude that both methods provide the same rank ordering, that confirm Oslo as the best candidate. It is noted that the two first alternatives, Oslo and Copenhagen, occupied the first position according to one indicator (Talent Rank and Green City Rank, respectively) and they were tied in the first position according to the Technology Rank. This apparently tie is broken in the final result taking into account, not only the input ordinal values, but also the dominance relations between pairs of alternatives along the considered indicators. On the other hand, Barcelona and Berlin are confirmed as the worst candidates among the studied cities.

V. CONCLUSIONS

A challenge of great interest in the context of smart cities is the problem of rank aggregation. Aggregate rankings, which compare city performance, are increasingly recognized as a useful tool in policy analysis, city marketing, benchmarking and public communication. The concept of smart city is a multi-criterial inherent character where different type of data related to subsets of cities, according different criteria and sources are supplied.

We have proposed two flexible rank-aggregation methods to produce aggregate rankings. Both summarize, in some precise sense, the information from input partial rankings. It addresses some of the shortcomings that appear in this context:

- The methods can handle partial lists: They do not require every pair of candidates (cities) x_i and x_j to be compared by every indicator. We use the available comparisons to construct the corresponding dominance matrix. The connectivity of the underlying graph information let infer information about pairs that have not been compared directly by any of the indicators, through the application of the method. Our methods allow incomparability, no data, and avoid the issue of assessing unknown data or to assume equal preference, as in some methods in the literature. Therefore, the proposed approach only takes into account available data. It improves the quality of the aggregate information in the sense that represents more fairly the input ranking evaluations.

- The assumption of minimum information required (the graph G is assumed simply connected or matrix A is irreducible) is less restrictive than other assumptions in the literature.

- The proposed methods learn preferences from the rankings lists, but do not treat the (associated) ordinal rank positions as if they were numerical values.

- Both methods are able to deal with lists that contain ties in their rankings.

- The computational formulations of the methods can be efficiently solved by numerical tools (least squares procedure, simplex method, and power method).

ACKNOWLEDGMENT

The authors are grateful to the referees for their valuable comments and suggestions that greatly improved the original manuscript.

REFERENCES

- [1] A. Abdoulaev. "A smart world: A development model for intelligent cities". *The 11th IEEE International Conference on Computer and Information Technology*, 2011, pp. 1-28.
- [2] S. Alonso, E. Herrera-Viedma, F. Chiclana, F. Herrera. "Individual and social strategies to deal with ignorance situations in multi-person decision making". *International Journal of Information Technology & Decision Making*, 8(02), pp. 313-333, 2009.
- [3] S. Brin, L. Page. "The anatomy of a large-scale hypertextual Web search engine". *Computer networks and ISDN systems*, 30(1), pp. 107-117, 1998.
- [4] H. Chourabi, T. Nam, S. Walker, J.R. Gil-Garcia, S. Mellouli, K. Nahon, H.J. Scholl. "Understanding smart cities: An integrative framework". *IEEE International Conference on System Science*, 2012, pp. 2289-2297.
- [5] W. D. Cook, L.M. Seiford. "Priority ranking and consensus formation". *Management Science*, 24(16), pp. 1721-1732, 1978.
- [6] Council. Smart Cities. "Smart cities readiness guide." 2013.
- [7] E. Dopazo, K. Lui, S. Chouinard, J. Guisse. "A parametric model for determining consensus priority vectors from fuzzy comparison matrices". *Fuzzy Sets and Systems*, 246, pp. 49-61, 2014.
- [8] E. Dopazo, M. Ruiz-Tagle. "A parametric GP model dealing with incomplete information for group decision-making". *Applied Mathematics and Computation*, 218, pp. 514-519, 2011.
- [9] C. Dwork, R. Kumar, M. Naor, D. Sivakumar. "Rank Aggregation Methods for the Web". *Proceedings of the ACM International Conference on World Wide Web*, 2001, pp. 613-622.
- [10] M. Farah, D. Vanderpooten. "An outranking approach for rank aggregation in information retrieval". *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2007, pp. 591-598.
- [11] M. S. Fox. "Foundation ontologies requirements for global city indicators". *Workshops at the Twenty-Eighth AAAI Conference on Artificial Intelligence*, 2014.[Online]. Available: <http://www.aaai.org/ocs/index.php/WS/AAAIW14/paper/download/8873/8261>
- [12] R. Giffinger, C. Fertner, H. Kramar, R. Kalasek, N. Pichler-Milanovic, E. Meijers. "Smart cities-Ranking of European medium-sized cities". Vienna University of Technology, 2007.[Online]. Available: http://curis.ku.dk/ws/files/37640170/smart_cities_final_report.pdf
- [13] J. González-Pachón, C. Romero. "Aggregation of partial ordinal rankings: an interval goal programming approach". *Computers & Operations Research*, 28(8), pp. 827-834, 2001.
- [14] E. Herrera-Viedma, F. Chiclana, F. Herrera, S. Alonso. "Group decision-making model with incomplete fuzzy preference relations based on additive consistency". *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 37(1), pp. 76-189, 2007.
- [15] E. Herrera-Viedma, F. Herrera, F. Chiclana, M. Luque. "Some issues on consistency of fuzzy preference relations". *European Journal of Operational Research*, 154, pp. 98-109, 2004.
- [16] D. S. Hochbaum, A. Levin. "Methodologies and algorithms for group-rankings decision". *Management Science*, 52(9), pp. 1394-1408, 2006.
- [17] J. P. Keener. "The Perron-Frobenius theorem and the ranking of football teams". *SIAM review*, 35(1), pp. 80-93, 1993.
- [18] J. G. Kemeny, L. J. Snell. "Preference ranking: an axiomatic approach". *Mathematical models in the social sciences*, pp. 9-23, 1962.
- [19] A. N. Langville, C. D. Meyer. *Who's #1?: The Science of Rating and Ranking*. Princeton University Press, 2012.
- [20] R. Martin, R. Florida, D. Tapscott. Martin Prosperity Institute (2013). Ranking Global Cities.[Online]. Available: http://www.cross-innovation.eu/wp-content/uploads/2013/11/Global-Cities-Insight_v01.pdf
- [21] P. McCarney. "Sustainable Development of Communities Indicators for City Services and Quality of Life". ISO 37120. International Organization for Standardization, 2015.

- [22] T. Moonen, G. Clark, R. Feenan. "The business of cities 2013: What do 150 city indexes and benchmarking studies tell us about the urban world in 2013". Available:<http://www.jll.com/Research/jll-city-indices-november-2013.pdf>
- [23] G. Munda. "Measuring sustainability: a multi-criterion framework". *Environment, Development and Sustainability*, 7(1), pp. 117-134, 2005.
- [24] G. Munda, M. Nardo. (2009). "Noncompensatory/nonlinear composite indicators for ranking countries: a defensible setting". *Applied Economics*, 41(12), pp. 1513-1523, 2009.
- [25] Renda, M. E., Straccia, U. "Web metasearch: rank vs. score based rank aggregation methods". *Proceedings of the 2003 ACM Symposium on Applied Computing*, 2003, pp. 841-84.
- [26] T.L. Saaty. *The Analytic Hierarchy Process*. Mc-Graw-Hill, New York, 1980.
- [27] T. Tanino. "Fuzzy preference orderings in group decision making". *Fuzzy Sets and Systems*, 12(2), pp. 117-131, 1984.
- [28] Unit, Economist Intelligence, and A. G. Siemens. "European Green City Index." Munich: Siemens AG 99 (2009). [Online]. Available: http://www.siemens.com/press/pool/de/events/corporate/2009-12-Cop15/European_Green_City_Index.pdf
- [29] R. Ureña, F. Chiclana, J. A. Morente-Molinera, E. Herrera-Viedma. "Managing Incomplete Preference Relations in Decision Making: A Review and Future Trends". *Information Sciences*, 302, pp. 14-32, 2015.